

MEASURING THE PERFORMANCE OF
FEDERAL AGENCIES AND PROGRAMS IN THE
USA: AN OVERVIEW AND SOME
REFLECTIONS

Abstract

In this paper, I provide an overview of some of the issues surrounding the measurement of government performance, focusing on the case of the U.S. federal government. After defining government performance and introducing the challenge of measuring it, I examine how this challenge is being addressed at the federal level in the United States. I argue that performance measurement is particularly difficult in this context because under the American constitution, agencies respond to multiple principals, in particular the president and Congress, which often have diverging preferences. I introduce congressional and presidential performance measurement initiatives, especially the Government Performance and Results Act including its recent reform, the Obama administration's high priority performance goals, and the George W. Bush administration's Program Assessment Rating Tool. I conclude with a brief review of potential undesired agency response to targets (such as the high priority performance goals), drawing on Christopher Hood's research on gaming.

I thank the Academy of Economic Studies, Faculty of Administration and Public Management for inviting me to present this paper as keynote speaker at its conference held in Bucharest, Romania, on June 21-22, 2011, and I thank Rob Greer for helpful research assistance.

Keywords: performance target, measuring performance, federal government.

JEL CODES: H11

**MĂSURAREA
PERFORMATEI
AGENȚIILOR ȘI
PROGRAMELOR
FEDERALE DIN SUA: O
ANALIZĂ ȘI CÂTEVA
REFLECȚII**

Nicolai PETROVSKY

Assistant Professor Ph.D., Martin School of
Public Policy & Administration, University of
Kentucky

E-mail: nicolai.petrovsky@uky.edu

Rezumat

În această lucrare voi prezenta o analiză a unor elemente conexe măsurării performanței guvernului, vizând situația guvernului federal din SUA. După definirea performanței guvernului și prezentarea provocărilor în măsurarea acestuia, voi examina cum această provocare este abordată la nivelul federal în Statele Unite. Argumentez că măsurarea performanței este dificilă în contextul constituției americane, agențiile fiind răspunzătoare în fața unor decidenți diferiți, precum Președintele și Congresul, care au de multe ori preferințe divergente. Voi prezenta inițiativele din Congres și prezidențiale de măsurare ale performanței, în special Legea Rezultatelor și Performanțelor Guvernamentale, incluzând reformele recente, obiectivele de performanță de importanță ridicată ale administrației Obama și Instrumentul de Clasificare a Evaluării Programelor al administrației Bush. Voi încheia cu o scurtă prezentare a potențialei reacții a agenției față de ținte (precum obiectivele de performanță de prioritate ridicată), pe baza unei cercetări a jocului, realizată de Christopher Hood.

Mulțumesc Academiei de Studii Economice-Facultății de Administrație și Management Public pentru ospitalitatea de a-mi adresa invitația să prezint această lucrare în calitate de keynote speaker la conferința pe care a organizat-o în București - România, în perioada 21-22 iunie 2011, și lui Rob Greer pentru utilul sprijin acordat în cercetare.

Cuvinte cheie: ținte de performanță, măsurarea performanței, guvern federal.



Proceedings of the seventh Administration
and Public Management International
Conference

1. INTRODUCTION

For some people, government performance is an oxymoron. It clearly does not have to be one though. What we can say with certainty is that government performance is a complex concept. The common denominator of a number of typically used definitions is that government performance is multidimensional (for a more detailed discussion, see Brewer and Selden, 2000, 688-689). For the purposes of this paper, I define government performance as the degree to which a government agency or program achieves the goals assigned to it through the democratic process. This definition is a narrowed-down variation of Rainey and Steinbauer's (1999) definition of agency effectiveness: "The agency performs well in discharging the administrative and operational functions pursuant to the mission. It achieves the mission as conceived by the organization and its stakeholders, or pursues achievement of it in an evidently successful way" (13). Government performance is a complex concept because not only does it always have more than one dimension but citizens disagree about the importance of the different dimensions. Worse, the different dimensions of performance may not even be correlated (for a related discussion, see Meyer, 2002, pp. 2-3). Boyne's (2002) breakdown of the different dimensions of local government performance is helpful as a general framework for understanding the concept. The fifteen dimensions of performance he enumerates clearly show the breadth of government performance: quantity of outputs, quality of outputs, cost per unit of output, formal effectiveness, impact, equity, cost per unit of formal effectiveness/impact/equity, consumer satisfaction, citizen satisfaction, staff satisfaction, cost per unit of consumer/citizen/staff satisfaction, probity, participation, accountability, and cost per unit of probity/participation/accountability (Boyne 2002, 19). It is easy to construct examples where some of these dimensions are negatively correlated or uncorrelated.

One important observation about government performance in all contexts is that its measurement is always imperfect. The city of Bristol in southern England provides a good example. As part of a central government mandate, Bristol, just like all other primary local governments, had been collecting random sample survey data on citizen satisfaction every three years. In the 2003 round of this survey, 64 per cent of respondents stated that they were very or fairly satisfied with the services provided by their local government. In the 2006 round, this figure dropped to a rather shocking 35 per cent. Had Bristol turned into a miserable failure in the meantime? Other data on Bristol's performance over this period do not suggest that there was a dramatic deterioration in its services. Rather, the most plausible explanation is a seemingly minor decision by Bristol Council: to switch from weekly to biweekly trash collection. The timing of this initially unpopular change---citizens now had to get used to keeping full trash bags on their

property for another week, in the warmest month of the year---was most unfortunate: It took place one month before the 2006 citizen satisfaction survey (Boyne et al. 2009; interview by the author and Oliver James with Bristol City Council performance officers on June 13, 2008). Citizen satisfaction can be even trickier as a measure since increases in performance may not match up to increased citizen expectations, such that an increase on a fairly clear measure of performance (such as the speed with which planning applications are processed) may be accompanied by a drop rather than by a rise in citizen satisfaction (James, 2009).

2. PERFORMANCE MEASUREMENT IN THE U.S. FEDERAL GOVERNMENT

While the measurement of government performance is far from easy at the local level, at the federal level in the United States the challenge takes on a whole other dimension. The peculiar constitutional context of the United States leads to further disagreement about what is to be measured, and how. Whereas in most European countries, a clear chain of delegation (Strøm, 2000, 266) extends from the electorate to parliament to the government to agencies, in the United States federal government, most agencies have more than one master. While they are formally under the ultimate command of the elected president, they are also under close scrutiny by the separately elected House of Representatives and the Senate, and especially the relevant subcommittees and appropriations committees in these two branches (Seidman, 1998). The incentives of agencies are generally in favor of closer alignment with Congress as opposed to the president. While a few scholars, particularly David Rosenbloom consider the president to essentially be an agent of Congress (Rosenbloom, 2000), presidents themselves and their electorates tend not to think or act in this fashion. Instead, they will do whatever is in their power to maintain and enhance their control over federal agencies. The result is an ongoing tug-of-war between president and Congress over the direction that federal agencies take (Macey, 1991; Whitford, 2005). In this tug-of-war, the major 'weapons' of Congress and its relevant committees and subcommittees are administrative rules (McCubbins, Noll and Weingast, 1989), informal signals sent to agencies including the threat of oversight (Seidman, 1998), incentives for interest groups to monitor agencies (McCubbins and Schwartz, 1984), and appropriations with conditions attached (Aldrich and Rohde, 2000; Copeland, 2008; Seidman, 1998). The president counters with appointments of agency heads (Epstein and O'Halloran, 1999; Wood and Waterman, 1991), signals sent to agencies in the presidential budget requests (Wood and Waterman, 1991), centralized control over all regulations issued by agencies in the Office of Management and Budget (Kagan, 2001; Seidman, 1998), and administrative reorganization, where feasible (Seidman, 1998). This interplay between the two elected branches is further complicated by the judiciary's interventions

(Macey, 1991), which are of minor relevance to the topic of the present paper, however. Moe and Caldwell (1994) provide an institutional explanation for this never-ending tug-of-war between Congress and president, and the complex nature of the American federal bureaucracy that results from it. Members of Congress tend to focus on specific geographic and sectoral constituencies. They tend to be most interested in systems that make agencies (even) more responsive to the particular subcommittees overseeing them. On the other hand, presidents tend to be interested in systems that allow them to centrally control and coordinate across agencies. One example of the latter is the Reagan administration's successful usage of the Carter administration's Paperwork Reduction Act to centralize control over all federal rule-making in the president's Office of Management and Budget (Seidman, 1998, 105-106). This system has served them well, and all presidents after Reagan have continued to use it.

Performance measurement in the U.S. federal government may be understood partly as a small component of the armories of Congress and the president in their respective attempts to steer agencies in the directions they desire.

The major weapon of Congress in this area has been the Government Performance and Results Act (GPRA). In its original form, it went into effect in 1993. There were three main aspects to the original GPRA, but they all revolved around the basic idea of federal agencies becoming more oriented towards results as opposed to process. First, federal agencies were required to develop a five-year strategic plan with mission statements and long term results-oriented goals for each of their major functions. Second, agencies were also required to develop annual performance plans that included performance goals, a description of how these goals would be met, and a method for measuring and evaluating these goals. Finally, agencies were required to prepare annual performance reports comparing actual results with the annual performance goals to assess success or failure in meeting targets. GPRA helped Congress to hold agencies accountable not only for processes but also results. The measurements prescribed by this legislation spurred a host of activity in federal performance measurement.

Seventeen years after the passage of the original GPRA, Congress passed the GPRA Modernization Act of 2010 (GPRAMA), and President Obama signed it into law on January 4, 2011. The new legislation takes into account the accumulated experience with the original GPRA. The major change lies in the greater integration of strategic plans, programs, and performance indicators in GPRAMA. The U.S. Government Accountability Office, Congress's watchdog agency, has identified five areas in which GPRAMA could address problem areas in government administration. First, GPRAMA could inform reexamination and restructuring efforts thanks to its more integrated approach. Second, it could

better address existing weaknesses in major management functions. Third, it could ensure that performance information is both useful and used in decision making processes. Fourth, it could achieve leadership commitment and accountability for results. Fifth, it could enhance Congress's engagement in identifying management and performance issues by increasing the requirements for agencies to consult with Congress (Dodaro, 2011). This last area is most pertinent to the thesis that GPRA and GPRAMA are helpful weapons in the congressional arsenal.

One of the most interesting changes brought by GPRAMA is that now, agencies' strategic plans are aligned with presidential terms rather than being written for a five-year period, as under the original GPRA. At the same time, the law strengthens congressional control by encouraging agencies to closely consult with Congress in the development of their strategic plans and performance goals. Overall, the act enhances the control of both elected branches in the area of performance management (see Kamensky, 2011).

GPRAMA also requires agencies to provide more information by revising the annual agency performance plan requirement. The performance plans must now include the strategies and resources needed to achieve the performance goals, and the plans must now extend for two years, rather than the previous one year periods. Not only are agencies required to specify clearly defined milestones, policies, regulations, and other activities necessary for achieving the performance goals, but they are also required to post their performance plan on the agency website (Kamensky, 2011).

Agencies have to submit their annual performance reports, which show how far along in achieving their performance plans they are, 150 days after the close of the fiscal year. The legislation encourages them to submit more frequent reports. GPRAMA brings about a major change to the performance reporting process by adding a "review and respond" requirement: Now, the president's Office of Management and Budget has to judge whether an agency is substantially achieving its performance goals. If the Office of Management and Budget determines that goals have not been met, they will submit a report to the head of the agency in question, the relevant congressional committees, and to the Government Accountability Office, the congressional watchdog agency. GPRAMA further stipulates a three-year plan for agencies that continue to fall short of their performance goals, with the reporting requirements becoming increasingly intrusive as the failure continues (Kamensky, 2011).

Perhaps the most interesting innovation of GPRAMA is that it requires the Director of the presidential Office of Management of Budget to work with agencies to develop federal government priority goals, which are policy goals that cut across agency boundaries (discussed further below). To promote the achievement of these policy goals, each agency head is required to identify which of their performance

goals align with the federal priority goals, and to prioritize these. The Director of the Office of Management of Budget will review the progress achieved toward these priority goals on a quarterly basis (Kamensky, 2011). This innovation means that GPRAMA enhances presidential control just as it enhances congressional control. It is important to note that GPRAMA was passed on a largely partisan basis towards the end of a period of Democratic Party control of both houses of Congress during the first part of Barack Obama's presidency (see House Republican Conference, 2010). This could mean that now, and in the future, when the preferences of Congress and its subcommittees once again diverge significantly from those of the president, a tension will develop between the congressional aspects of GPRAMA and the presidential aspects, in particular the high priority performance goals.

The Obama administration's high priority performance goals were introduced in the presidential budget request for fiscal year 2011 (October 1, 2010-September 30, 2011). Agency heads chose goals that they anticipated being able to achieve within two years and that did not require additional resources or legislative action. For example, the Social Security Administration specified four priority goals. The first is to increase the proportion of benefits applications filed online to 50 percent by 2012. The second is to lower the backlog in the processing of disability benefit claims such as to ensure that people filing a severely disabling conditions claim receive an initial decision within 20 days of filing their application. The third is to increase several overall public satisfaction metrics. The fourth is to minimize improper payments and strengthen efforts to protect program dollars from waste, fraud, and abuse (White House, 2011). Only the first two goals are clearly operational, i.e. it will be unambiguously clear at their stated end time whether or not they have been achieved (Wilson, 1989, 34).

Another example is the Department of Defense, which has identified eight categories of goals in everything from energy efficiencies to their acquisition process. The priority goals include producing or procuring renewable energy equal to 10 percent of the Department of Defense's annual electric energy usage, reforming the Personnel Security Clearance Process such that the fastest 90 percent of initial top secret and secret personnel security clearance cases are adjudicated within 20 days, and decreasing the Defense Department's reliance on contract services in acquisition functions by increasing the in-house workforce by 4765 authorizations for personnel (White House, 2011).

In a final example, among the Department of Housing and Urban Development's priority goals is a collaborative effort with the Department of Veteran Affairs to reduce the number of homeless veterans to 59,000 in June 2012 (White House, 2011).

Before High Priority Performance Goals, the previous performance initiative from the executive side was the Program Assessment Rating Tool (PART). It was the George W. Bush administration's approach to

enhancing the performance reporting for federal programs. Just as the current priority performance goals are overseen by the president's Office of Management and Budget, it also directed the PART program. Beginning in 2002, Office of Management and Budget staff reviewed scores of federal programs every year, up to and including 2008. The PART was a questionnaire with various items on each of four dimensions: (i) program purpose and design, (ii) planning, (iii) management, and (iv) results. Office of Management and Budget grading teams filled in answers to these questions in accordance with responses and evidence provided by program officials from the various federal agencies. The answers to the questions on each of the four dimensions were aggregated into numerical scores as well as an overall program rating, which was either an overall program performance evaluation ("effective," "moderately effective," "adequate," "ineffective") or an indication that the program did not provide the Office of Management and Budget with enough evidence for it to make a well-substantiated evaluation of the results of the program ("results not demonstrated"). The Office of Management and Budget published the aggregated PART scores for each of the four dimensions and the overall PART program rating as part of the presidential budget requests during the years of PART's existence. The overall ratings were also disseminated to the public through the website www.expectmore.gov. The Obama administration terminated PART in 2010.

3. INHERENT PROBLEMS IN THE USAGE OF PERFORMANCE TARGETS

At the end of this brief overview of performance measurement in the American federal government, it is worth taking a step back to consider some problems inherent in the usage of performance targets, such as the ones the Obama administration is now using and the British government has been using for over a decade. Drawing on academic research on the Soviet Union's usage of performance targets, Hood (2006) and Bevan and Hood (2006) carefully dissect the gaming behaviors that are virtually inevitable if performance targets are imposed. Among these behaviors, Hood (2006) distinguishes between ratchet, threshold, and output distortion effects.

Ratchet effects can occur whenever next year's performance targets are set as increments of this year's performance. If managers and employees know this process, they are likely to underachieve this year, so that next year's performance target, which will be an increase on this year's actual achievement, will not be challenging to them. It would be conceivable that a target of 50 percent online applications for Social Security benefits could be perceived as only the beginning of a series of successively more ambitious targets. If that were the case, while 60 percent might be in reach, the agency would limit its efforts toward that end to avoid being saddled with a truly difficult target such as 70 percent.

Threshold effects can occur when uniform performance targets are applied to a number of agencies or divisions, where some perform higher at the outset than others. The divisions performing highly from the outset have no incentive to improve since they already perform beyond the threshold specified by the target. Indeed, they may reduce their efforts and objectively fall in their performance, to the exact level of the target. Looking at the aforementioned examples of high priority performance goals, it would be conceivable that some security clearances in the Department of Defense and some Social Security disability benefit claims may now take longer than previously---e.g. some that could be done in 10 days might now take close to 20---as these agencies might shift resources in order to achieve the high priority performance goal of having security clearances and disability benefit claims completed within 20 days.

Output distortion effects are present whenever managers and employees focus on the letter of a performance target instead of its spirit. There appears to be potential for such an effect in the goal to reduce the number of homeless veterans to 59,000. The Department of Housing and Urban Development and the Department of Veteran Affairs might end up finding some solution that strictly fulfills the target by the deadline but fails to offer a longer-term perspective to the veterans affect, such that the number of homeless or dependent veterans might not decrease in a sustained manner.

I would like to stress that the examples of potential gaming discussed here are purely hypothetical, and that it is likely that Office of Management and Budget officials will take them into account in collecting information about agency progress toward these targets. These hypothetical examples do, however, serve the purpose of highlighting that performance management using targets is a challenging endeavor. It may be the case that it works for bringing agencies and programs up to some standard of performance, but that they are incapable of inducing excellence, as the dysfunctional side-effects of performance measurement become more important at higher baseline levels of performance. Future work will have to examine this hypothesis.

4. CONCLUSIONS

Regardless of future revisions to the recently reformed GPRA, or perhaps it being moved to the sideline by new initiatives, performance measurement and management will remain a salient topic at the American federal level. One reason is that critical citizens are likely to provide a continued impetus for efforts to improve the performance of government agencies and programs. In addition, a government that continues to raise less than 60 percent of its expenditures in taxes (United States Department of the Treasury Financial Management Service, 2011) may even find performance information to be one among several helpful inputs in making adjustments. The Obama administration has recently opened a

new web portal to make a number of metrics accessible to a broader public: <http://www.performance.gov/>. It will remain to be seen whether this leads to greater citizen involvement in appraising federal agency performance.

In spite of the virtually assured continued interest in performance measurement and management at the federal level in the United States, the major caveat is that it is, and will remain, of secondary importance in shaping the directions of federal agencies and programs compared to the tug-of-war between the president and the balance of interests in Congress.

REFERENCES

- Aldrich, J.H. and Rohde, D.W. (2000). The Republican Revolution and the House Appropriations Committee. *Journal of Politics* 62 (1), pp. 1-33.
- Bevan, G. and Hood, C. (2006). What's Measured Is What Matters: Targets and Gaming in the English Public Health Care System. *Public Administration* 84 (3), pp. 517-538.
- Boyne, G.A. (2002). Concepts and Indicators of Local Authority Performance: An Evaluation of the Statutory Frameworks in England and Wales. *Public Money & Management* 22 (2), pp. 17-24.
- Boyne, G.A., James, O., John, P. and Petrovsky, N. (2009). Leadership Change and Public Service Performance in English Local Governments, 1998-2006. Data set. Retrieved September 6, 2011, from <http://store.data-archive.ac.uk/store/>.
- Brewer, G.A. and Selden, S.C. (2000). Why Elephants Gallop: Assessing and Predicting Organizational Performance in Federal Agencies. *Journal of Public Administration Research and Theory* 10 (4), pp. 685-712.
- Copeland, C.W. (2008). *Congressional Influence on Rulemaking and Regulation Through Appropriations Restrictions*. Congressional Research Service Report for Congress, Order Code RL34354. Washington, DC: Congressional Research Service.
- Dodaro, G.L. (2011). General Audit Office Testimony: Managing for Results: GPRA Modernization Act Implementation Provides Important Opportunities to Address Government Challenges. Retrieved August 10, 2011, from <http://www.gao.gov/new.items/d11617t.pdf>.
- Epstein, D. and O'Halloran, S. (1999). *Delegating Powers: A Transaction Cost Politics Approach to Policy Making under Separate Powers*. Cambridge: Cambridge University Press.
- Hood, C. (2006). Gaming in Targetworld: The Targets Approach to Managing British Public Services, *Public Administration Review* 66 (4), pp. 515 - 521.
- House Republican Conference. (2010). *Roll Call 660 on H.R. 2142: On Motion To Concur In The Senate Agreement*. Retrieved June 17, 2011 from <http://www.gop.gov/votes/111/2/660>.
- James, O. (2009). Evaluating the Expectations Disconfirmation and Expectations Anchoring Approaches to Citizen Satisfaction with Local Public Services. *Journal of Public Administration Research and Theory* 19 (1), pp. 107-123.
- Kagan, E. (2001). Presidential Administration. *Harvard Law Review* 114 (8), pp. 2245-2385.

- Kamensky, J.M. (2011). *GPRA Modernization Act of 2010*. IBM Center for The Business of Government. Retrieved August 10, 2011, from <http://www.scribd.com/doc/47464749/GPRA-Modernization-Act-of-2010-Explained>.
- Macey, J.R. (1991). Separated Powers and Positive Political Theory: The Tug of War Over Administrative Agencies. *Georgetown Law Journal* 80, pp. 671-703.
- McCubbins, M.D. and Schwartz, T. (1984). Congressional Oversight Overlooked: Police Patrols versus Fire Alarms. *American Journal of Political Science* 28 (1), pp. 165-179.
- McCubbins, M.D., Noll, R.G. and Weingast, B.R. (1989). Structure and Process, Politics and Policy: Administrative Arrangements and the Political Control of Agencies. *Virginia Law Review* 75 (2), pp. 431-482.
- Meyer, M.W. (2002). *Rethinking Performance Measurement: Beyond the Balanced Scorecard*. Cambridge: Cambridge University Press.
- Moe, T.M. and Caldwell, M. (1994). The Institutional Foundations of Democratic Government: A Comparison of Presidential and Parliamentary Systems. *Journal of Institutional and Theoretical Economics* 150 (1), pp. 171-195.
- Rainey, H.G. and Steinbauer, P. (1999). Galloping Elephants: Developing Elements of a Theory of Effective Government Organizations. *Journal of Public Administration Research and Theory* 9 (1), pp. 1-32.
- Rosenbloom, D.H. (2000). *Building a Legislative-Centered Public Administration: Congress and the Administrative State, 1946-1999*. Tuscaloosa: University of Alabama Press.
- Seidman, H. (1998). *Politics, Position, and Power: The Dynamics of Federal Organization*. 5th edition. Oxford: Oxford University Press.
- Strøm, K. (2000). Delegation and Accountability in Parliamentary Democracies. *European Journal of Political Research* 37 (3), pp. 261-290.
- United States Department of the Treasury Financial Management Service. (2011). *Monthly Receipts, Outlays, and Deficit or Surplus, Fiscal Years 1981-2010*. Retrieved September 16, 2011, from <http://www.fms.treas.gov/mts/mts.pdf>.
- White House. (2011). *Executive Office of the President of the United States High Priority Performance Goals*. Retrieved August 10, 2011, from <http://www.whitehouse.gov/sites/default/files/omb/performance/high-priority-performance-goals.pdf>.
- Whitford, A.B. (2005). The Pursuit of Political Control by Multiple Principals. *Journal of Politics* 67 (1), pp. 28-49.
- Wilson, J.Q. (1989). *Bureaucracy: What Government Agencies Do and Why They Do It*. New York: Basic Books.
- Wood, B.D. and Waterman, R.W. (1991). The Dynamics of Political Control of the Bureaucracy. *The American Political Science Review* 85 (3), pp. 801-828.